

21-CM TOMOGRAPHY OF THE INTERGALACTIC MEDIUM AT HIGH REDSHIFT

Piero Madau

Space Telescope Science Institute, 3700 San Martin Drive, Baltimore MD 21218; madau@stsci.edu

Avery Meiksin¹

Dept. of Astronomy and Astrophysics, University of Chicago, 5640 S. Ellis Ave., Chicago IL

60637; meiksin@oddjob.uchicago.edu

and

Martin J. Rees

Institute of Astronomy, Madingley Road, Cambridge CB3 0HA

Received _____; accepted _____

¹Edwin Hubble Research Scientist

ABSTRACT

We investigate the 21-cm signature that may arise from the intergalactic medium (IGM) prior to the epoch of full reionization ($z > 5$). In scenarios in which the IGM is reionized by discrete sources of photoionizing radiation, the neutral gas which has not yet been engulfed by an H II region may easily be preheated to temperatures well above that of the cosmic background radiation (CBR), rendering the IGM invisible in absorption against the CBR. We identify three possible preheating mechanisms: (1) photoelectric heating by soft X-rays from QSOs, (2) photoelectric heating by soft X-rays from early galactic halos, and (3) resonant scattering of the continuum UV radiation from an early generation of stars. We find that bright quasars with only a small fraction of the observed comoving density at $z \sim 4$ will suffice to preheat the entire universe at $z \gtrsim 6$. We also show that, in a Cold Dark Matter dominated cosmology, the thermal bremsstrahlung radiation associated with collapsing galactic mass halos ($10^{10} - 10^{11} M_{\odot}$), may warm the IGM to ~ 100 K by $z \sim 7$. Alternatively, the equivalent of $\sim 10\%$ of the star formation rate density in the local universe, whether in isolated pregalactic stars, dwarf, or normal galaxies, would be capable of heating the entire IGM to a temperature above that of the CBR by Ly α scattering in a small fraction of the Hubble time at $z \sim 6$.

In the presence of a sufficiently strong ambient flux of Ly α photons, the hyperfine transition in the warmed H I will be excited. A beam differencing experiment would detect a patchwork of *emission*, both in frequency and in angle across the sky. This patchwork could serve as a valuable tool for understanding the epoch, nature, and sources of the reionization of the universe, and their implications for cosmology. We demonstrate that isolated QSOs will produce detectable signals at meter wavelengths within their “spheres of influence” over which they warm the IGM. As a result of the redshifted 21-cm radiation emitted by warm H I bubbles, the spectrum of the radio extragalactic background will display frequency structure with velocity widths

up to $10,000 \text{ km s}^{-1}$. Broad beam observations would reveal corresponding angular fluctuations in the sky intensity with $\delta T/T \lesssim 10^{-3}$ on scales $\theta \sim 1^\circ$. This scale is set either by the “thermalization distance” from a QSO within which $\text{Ly}\alpha$ pumping determines the spin temperature of the IGM or by the quasar lifetime.

Radio measurements near 235 and 150 MHz, as will be possible in the near future using the *Giant Metrewave Radio Telescope*, may provide the first detection of a neutral IGM at $5 \lesssim z \lesssim 10$. A next generation facility like the *Square Kilometer Array Interferometer* could effectively open much of the universe to a direct study of the reheating epoch, and possibly probe the transition from a neutral universe to one that is fully ionized.

Subject headings: cosmology: theory – diffuse radiation – intergalactic medium – quasars: general

1. Introduction

At epochs corresponding to $z \sim 1000$, the intergalactic medium (IGM) is expected to recombine and remain neutral until sources of radiation develop that are capable of reionizing it. The application of the Gunn-Peterson (1965) constraint on the amount of cosmologically distributed neutral hydrogen to QSO absorption spectra requires the universe to have been highly ionized by $z \sim 5$ (Schneider, Schmidt, & Gunn 1991). It thus appears that substantial sources of ionization were present at these or earlier epochs, perhaps QSOs and young galaxies, or some as yet undiscovered class of objects such as Pop III stars or decaying particles. Establishing the epoch of reionization is crucial for determining the impact of reionization on several key cosmological issues, from the role reionization plays in allowing collapsed objects to cool and make stars, to determining the small-scale structure in the temperature fluctuations of the cosmic background radiation (CBR) (e.g., Sugiyama, Silk, & Vittorio 1993; Bond 1995; Hu & White 1996). Conversely, probing the epoch of reionization may provide a means of detecting the onset of the first generation of stars.

Very little is known about the nature of the first bound objects and of the thermal state of the universe at early epochs. In particular, the history of the transition from a neutral IGM to an almost fully ionized one can reveal the character of the ionizing sources. If these are uniformly distributed, like an abundant population of pregalactic stars or decaying dark matter, the ionization and thermal state of the IGM will be the same everywhere at any given epoch, with the neutral fraction decreasing gradually with cosmic time. In a CDM dominated cosmology, this may be the case since bound objects sufficiently massive ($\sim 10^6 M_\odot$) to make stars form at high redshift. If a sufficient fraction of their ionizing photons are able to escape, the stars in these objects would reionize the universe by $z \gtrsim 20$ (Couchman & Rees 1986). In inhomogeneous reionization scenarios instead, widely separated but very luminous sources of photoionizing radiation such as QSOs, present at the time the IGM is largely neutral, will generate expanding H II regions whose size is only limited by the individual source lifetime. The universe will be divided into an ionized phase whose filling factor $q(z)$ increases with time, and an ever shrinking

neutral phase (Arons & Wingert 1972; Meiksin & Madau 1993). If the ionizing sources are randomly distributed, the H II regions will be spatially isolated for $q \ll 1$. The IGM will become completely reionized at the breakthrough epoch, when $q = 1$ and the IGM is fully transparent at the Lyman continuum. Current QSO counts suggest that the breakthrough epoch may occur as recently as $z \gtrsim 5$ (Meiksin & Madau 1993).

It is therefore of great interest, in testing rival reionization scenarios, to seek any evidence of non-uniformity in the ionization structure of the IGM by studying the thermal and ionization history of the intergalactic gas surrounding isolated sources of radiation (see, e.g., Miralda-Escudè & Rees 1994). In this paper we shall discuss how radio-astronomical measurements of spectral and spatial structures in the 21-cm line emission and absorption from H I at high redshifts may provide a valuable tool for probing the epoch of the likely heat input at $5 \lesssim z \lesssim 10$ and its nature, e.g., whether it was limited to small isolated regions or was relatively diffuse. We will confine our analysis to epochs prior to complete reionization, and address the physical state of the gas which has not yet been engulfed by an ionized region.

For intergalactic H I to be observable in 21-cm line emission or absorption against the cosmic background radiation, the spin temperature must differ from the CBR temperature. As the former will in general be a weighted mean between the matter and CBR temperatures, an efficient mechanism for coupling the populations of the hyperfine levels to the kinetic state of the intergalactic gas is needed to unlock the spin state from the CBR. At the low particle densities characteristic of the IGM, the dominant coupling mechanism is level-mixing via Ly α photon scattering – the “Wouthuysen-Field” effect. We present in this paper a new heating mechanism, arising from atomic recoil in response to the scattering of resonant photons. We demonstrate that because of resonant scattering heating, Ly α coupling makes the detection of the IGM in absorption against the CBR quite unlikely, since the same Ly α photons that pump the hyperfine levels heat the atoms as well. This is not entirely surprising, as the energetic demand for raising the kinetic temperature above that of the CBR is relatively small, $\sim 10^{-7} \text{ eV cm}^{-3}$ at $z \sim 7$, hence even a relatively inefficient heating mechanism like Ly α resonant scattering may become important.

We find that the heating requirements are quite generally met in a variety of scenarios in which radiation sources turn on at high redshifts. We show that, in models based on QSO photoionization, and well before the H II region network has fully percolated, the (mostly) neutral IGM between the H II bubbles will be photoelectrically heated to temperatures above a few hundred degrees by soft X-ray photons from the QSOs. In Cold Dark Matter (CDM) models a similar mechanism may efficiently warm the IGM at $z \lesssim 7$, since galactic mass halos with virial temperatures $\sim 10^6$ K are significant sources of thermal bremsstrahlung radiation. Alternatively, if protogalaxies at these early epochs are forming stars at a rate which is only $\sim 10\%$ of the global present-day star formation rate, heating by background Ly α “continuum” photons could drive the kinetic temperature of the still neutral material well above the radiation temperature of the microwave background. As a consequence, we conclude that *most of the neutral IGM will be available for detection at 21-cm only in emission.*

Two scenarios serve to illustrate the observability at meter wavelengths of the intergalactic medium during reionization. An isolated QSO turning on in a neutral medium will create an H II bubble surrounded by an X-ray warmed neutral zone. Sufficiently near the QSO, the spin state of the neutral IGM will be coupled to its kinetic state by redshifted continuum Ly α photons from the QSO. This zone will then emit 21-cm radiation that would be detectable against the CBR as a differential intensity, in angle or frequency. A stronger signal may be achieved in the presence of a Ly α continuum background sufficiently strong to couple the spin temperature of the cold and neutral IGM outside the QSO “sphere of influence” to the kinetic temperature of the IGM. The larger signal results from the enhanced contrast between the 21-cm absorbing medium far from the QSO, and the 21-cm emitting gas in the X-ray warmed neutral zone surrounding the QSO H II region. This phase, however, will be short-lived: through atomic recoil the Ly α background photons will soon heat the IGM far from the QSO to a temperature above that of the CBR, and absorption will cease.

The plan of the paper is as follows. In § 2 we discuss sources of radiation at high redshift. In § 3 we review the mechanisms that determine the level populations of the hyperfine state of

neutral hydrogen. In § 4 we treat the preheating of the IGM prior to complete reionization, and show that a 21-cm absorbing IGM can exist only as a transitory state. In § 5 we argue that the 21-cm line signature from a neutral IGM, redshifted to meter wavelengths, will appear as a small feature in the radio extragalactic background on scales of $\lesssim 1^\circ$, and may be detectable with the *Giant Metrewave Radio Telescope*. Finally, in § 6, we summarize our conclusions.

2. Sources of Radiation at High Redshift

In the following, we assume that the universe is completely reionized by discrete sources at $z \gtrsim 5$, and that the IGM at such early epochs is uniform on large scales. Unless otherwise stated, we shall adopt a flat cosmology with $q_0 = 0.5$ and $H_0 = 50 h_{50} \text{ km s}^{-1} \text{ Mpc}^{-1}$.

2.1. QSO H II Regions

When an isolated point source of ionizing radiation turns on, the ionized volume initially grows in size at a rate fixed by the emission of UV photons, and an ionization front separating the ionized and neutral regions propagates into the neutral gas. The evolution of an expanding cosmological I-front in a uniform IGM is governed by the equation

$$4\pi r_I^2 n_H \left(\frac{dr_I}{dt} - H r_I \right) = S - \frac{4}{3} \pi r_I^3 (1 + 2\chi) n_H^2 \alpha_B, \quad (1)$$

(Shapiro 1986), where r_I is the proper radius of the I-front, n_H is the hydrogen density of the IGM, H is the Hubble constant, S is the number of ionizing photons emitted by the central source per unit time, $\alpha_B \simeq 2.6 \times 10^{-13} T_4^{-0.845} \text{ cm}^3 \text{ s}^{-1}$ is the recombination coefficient to the excited states of hydrogen, and T_4 is the gas temperature in units of 10^4 K . A helium to hydrogen cosmic abundance ratio $\chi = 1/12$ is adopted, with both H and He fully ionized inside the H II region. The right hand side of equation (1) is a measure of the net flux of ionizing photons reaching the I-front after subtracting recombinations, and vanishes at the Strömgren radius, $r_{\text{St}} = \{3S/[4\pi n_H^2 \alpha_B (1 + \chi)]\}^{1/3}$. Across the I-front the degree of ionization changes sharply on a distance of the order of the mean

free path of an ionizing photon. Denoting by Ω_{IGM} the baryonic mass density parameter of the intergalactic medium, the I-front will expand at the speed of light

$$r_I < \left(\frac{S}{4\pi n_{\text{H}} c} \right)^{1/2} \approx (2.85 \text{ Mpc}) \left(\frac{S}{10^{57} \text{ s}^{-1}} \right)^{1/2} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^{-1/2} \left(\frac{1+z}{7} \right)^{-3/2}, \quad (2)$$

to subsequently slow down because of geometrical dilution. In an expanding IGM, the ionization front generated by a QSO, with $S \sim 10^{57} \text{ s}^{-1}$, or a star-forming galaxy, with $S \sim 10^{53} \text{ s}^{-1}$, will fail to grow to even half of its Strömngren radius (Shapiro 1986; Madau & Meiksin 1991). In this case radiative recombinations within the H II region may be neglected on the right-hand side of equation (1) (proportional to $r_{\text{St}}^3 - r_I^3$), to obtain at a short time after the turn-on epoch z (corresponding to the redshift interval $\Delta z \ll 1+z$),

$$r_I \approx (10.2 \text{ Mpc}) \left(\frac{S}{10^{57} \text{ s}^{-1}} \right)^{1/3} h_{50}^{-1/3} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^{-1/3} \left(\frac{1+z}{7} \right)^{-11/6} \Delta z^{1/3}, \quad (3)$$

and

$$v_I \approx (2.2 \times 10^4 \text{ km s}^{-1}) \left(\frac{S}{10^{57} \text{ s}^{-1}} \right)^{1/3} h_{50}^{2/3} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^{-1/3} \left(\frac{1+z}{7} \right)^{2/3} \Delta z^{-2/3}. \quad (4)$$

For comparison, the radius of the light front around each source is

$$r_c \equiv c\Delta t \approx (46.3 \text{ Mpc}) h_{50}^{-1} \left(\frac{1+z}{7} \right)^{-5/2} \Delta z. \quad (5)$$

2.2. Virialized Halos

In CDM dominated cosmologies, galactic mass halos will begin to collapse at high redshift ($z > 5$). The hot gas in these halos will produce a soft X-ray background radiation field via thermal bremsstrahlung. We estimate the number of halos using the Press-Schechter formalism (Press & Schechter 1974), which should suffice for an order of magnitude estimate for systems on these scales.

The (proper) volume emissivity per unit frequency due to thermal bremsstrahlung radiation from isothermal, optically thin halos is

$$\epsilon_{\text{halo}}(\nu) = (3.2 \times 10^{36} \text{ ergs Mpc}^{-3} \text{ s}^{-1} \text{ Hz}^{-1}) \int_0^\infty dr_0 N(r_0, z) (1+z)^3 n_{\text{halo}}^2(r_0) \frac{4\pi}{3} r_v^3 T(r_0)^{-1/2} \times \exp[-h\nu/kT(r_0)], \quad (6)$$

where $N(r_0, z)$ is the number of halos per unit comoving volume per initial comoving size r_0 at redshift z , $n_{\text{halo}}(r_0)$ is the internal hydrogen density of a halo, and $T(r_0)$ is its temperature. The hydrogen is assumed to be shock heated to the virial temperature and collisionally ionized. In spherical top-hat collapse, the temperature T and radius r_v of a virialized halo are related to r_0 by $kT(r_0) = 1.4\mu m_{\text{H}}(1+z_c)H_0^2 r_0^2$ and $r_0 = 5.6(1+z_c)r_v$, where z_c is the collapse redshift of the perturbation, and $\mu \approx 0.6$ is the mean molecular weight of a fully ionized gas of cosmic abundances. The internal density of the cloud exceeds the external density at the projected time of collapse by a factor of 178. Normalizing to a baryon density of $\Omega_{\text{IGM}}h_{50}^2 = 0.05$ gives $n_{\text{halo}} = 1.8 \times 10^{-5}(1+z_c)^3 \text{ cm}^{-3}$. According to the Press-Schechter formalism,

$$N(r_0, z)dr_0 = -3 \frac{1.686(1+z)}{(2\pi)^{3/2}r_0^4\Delta(r_0)} \frac{d\log\Delta(r_0)}{d\log(r_0)} \exp\left[-\frac{1.686^2(1+z)^2}{2\Delta(r_0)^2}\right] dr_0. \quad (7)$$

Here, $\Delta(r_0)$ is the rms mass fluctuation in a sphere of radius r_0 . We will adopt a pure CDM spectrum for a $q_0 = 0.5$, $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ universe (Bardeen et al. 1986). This model appears to provide an adequate representation of the power spectrum on small scales when appropriately renormalized, even though a higher normalization is needed for an optimal match to the temperature fluctuations measured by COBE on large scales (Górski et al. 1995). Normalizing to the abundance of clusters, $\Delta(16 \text{ Mpc}) \approx 0.6$ (White et al. 1993), we find that for $0.2 < r_0 < 1 \text{ Mpc}$, $\Delta(r_0) \approx 3.6r_0^{-3/8}$, where r_0 is measured in megaparsecs.

The total emissivity above a minimum photon energy $h\nu_{\text{min}}$ is $\epsilon_{\text{halo}} = [kT(r_0)/h]\epsilon_{\text{halo}}(\nu_{\text{min}})$, after integrating equation (6) over frequency. While the emissivity from an individual halo increases for increasing temperature (for $kT < h\nu_{\text{min}}$), the number of halos decreases rapidly. At any given redshift, the peak contribution to the integral in equation (6) arises therefore from some intermediate size halo. We may thus approximate this integral by performing a second order Taylor expansion of the exponent about the halo size. Most of the bremsstrahlung radiation is found to be emitted by halos with an initial comoving radius

$$\hat{r}_0 \approx (1 \text{ Mpc}) \left(\frac{h\nu_{\text{min}}}{20 \text{ ryd}}\right)^{4/11} \left(\frac{1+z}{7}\right)^{-12/11}. \quad (8)$$

The total emissivity is given by

$$\epsilon_{\text{halo}} \approx (2.6 \times 10^{43} \text{ ergs Mpc}^{-3} \text{ s}^{-1}) \hat{r}_0^{3/8} \left(\frac{1+z}{7} \right)^7 \left[\frac{kT(\hat{r}_0)}{20 \text{ ryd}} \right] \left(\frac{h\nu_{\text{min}}}{20 \text{ ryd}} \right)^{-1/2} \times \exp \left[-5.3 \left(\frac{1+z}{7} \right)^2 \hat{r}_0^{3/4} - 1.8 \left(\frac{1+z}{7} \right)^{-1} \left(\frac{h\nu_{\text{min}}}{20 \text{ ryd}} \right) \hat{r}_0^{-2} \right]. \quad (9)$$

We shall use this expression in § 4.2.3 below to estimate the virialized halo contribution to the heating rate of the IGM.

2.3. Early Stellar Populations

An early generation of stars may provide a background radiation field of Ly α photons that arises from the collective redshifted continua of the stars. The background specific intensity associated with a constant comoving emissivity from the present epoch to redshift z_{max} , as seen by an observer at redshift z , is given by

$$J_{\alpha}(z) = \frac{1}{4\pi} \epsilon(\nu_{\alpha}, 0) (1+z)^{3+\alpha_S} \int_z^{z_{\text{max}}} \frac{d\ell}{dz'} (1+z')^{-\alpha_S} dz', \quad (10)$$

where $\epsilon(\nu_{\alpha}, 0)$ is the Ly α volume emissivity due to stars in galaxies at the present epoch, $d\ell/dz'$ is the proper differential line element in a Friedmann cosmology, and we have assumed a power law energy spectrum. Here, $1+z_{\text{max}} \approx 4(1+z)/3$, as photons emitted at higher redshifts will have wavelength below 912 Å at the source, and will be strongly attenuated by Lyman-continuum intergalactic absorption. Weakly dependent on α_S , equation (10) may be rewritten as

$$J_{\alpha}(z) \approx (10^{-20.7} \text{ ergs cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1}) h_{50}^{-1} \left[\frac{\epsilon(\nu_{\alpha}, 0)}{10^{25} \text{ ergs Mpc}^{-3} \text{ s}^{-1} \text{ Hz}^{-1}} \right] \left(\frac{1+z}{7} \right)^{3/2}. \quad (11)$$

Using the results of the H α survey of Gallego et al. (1995), we estimate the continuum emissivity from present day galaxies at the frequency of Ly α to be, to within a factor of 2, $\epsilon(\nu_{\alpha}, 0) \approx 4 \times 10^{25} \text{ ergs Mpc}^{-3} \text{ s}^{-1} \text{ Hz}^{-1}$. Only a fraction of this emissivity would be available at high redshift from dwarf galaxies or the progenitors of today's bright galaxies. The Ly α continuum emissivity at early epochs, however, may be even larger if pregalactic Pop III stars are present at these times.

The IGM will still be mostly neutral if the background intensity at the Lyman edge, J_L , satisfies

$$J_L < \frac{hc}{4\pi} n_{\text{H}}(0)(1+z)^3 \approx (10^{-21.2} \text{ ergs cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1}) \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right) \left(\frac{1+z}{7} \right)^3, \quad (12)$$

i.e., if less than one ionizing photon per hydrogen atom has reached the IGM by that epoch. Because the integrated UV spectrum of a stellar population is characterized by a strong Lyman-continuum break (about a factor of 4), and, more importantly, only a small fraction of ionizing photons might escape from collapsed star-forming regions into intergalactic space (cf. Madau & Shull 1996), one expects $J_L \ll J_\alpha$. Hence, based on a comparison of equations (11) and (12), it is plausible to assume a largely neutral medium even in the presence of a significant Ly α background from an early generation of stars, $J_{\alpha,-21} \approx 1$.

3. Spin Temperature of the IGM

The amount of emission or absorption from a neutral intergalactic medium is determined by the spin state of the atoms. The spin temperature may be coupled to the matter both through the scattering of Ly α photons – the Wouthuysen-Field effect (Wouthuysen 1952; Field 1958) – and atomic collisions. At the low particle densities characteristic of intergalactic gas, and in the presence of radio-quiet radiation sources, the dominant coupling mechanism is Ly α scattering.

3.1. Ly α Radiation Color Temperature

The Wouthuysen-Field effect mixes the hyperfine levels of neutral hydrogen via the intermediate step of transitions to the $2p$ state. According to this mechanism, an atom initially in the $n = 1$ $^1S_{1/2}$ singlet state may absorb a Ly α photon that will put it in the $n = 2$ $^1P_{1/2}$ or $^1P_{3/2}$ state, allowing it to return to the triplet $n = 1$ $^3S_{1/2}$ state by spontaneous decay. Similarly, a triplet atom may absorb a slightly lower frequency photon to reach the same upper state, followed by decay to the singlet state. The efficiency of the effect is determined by the frequency

dependence of the background radiation field J_ν near $\text{Ly}\alpha$. In particular, the excitation rates of the hyperfine levels depend on the relative intensities of the blue and the red wings of $\text{Ly}\alpha$. To observe the 21-cm line in absorption or emission, the excitation and de-excitation rates of the hyperfine levels must be comparable to or exceed those due to 21-cm continuum background radiation.

The $\text{Ly}\alpha$ line absorption cross section is

$$\sigma_\nu = \sigma_\alpha \phi(\nu) = \frac{\pi e^2}{m_e c} f_\alpha \phi(\nu) = \frac{3}{8\pi} \lambda_\alpha^2 A_\alpha \phi(\nu) \quad (13)$$

(neglecting the very small correction due to stimulated emission, of order $\lambda_\alpha^3 J_\alpha / 2hc \ll 1$), where A_α and f_α are the spontaneous Einstein coefficient and upward oscillator strength for the transition, $\phi(\nu)$ is the line profile function (with normalization $\int \phi(\nu) d\nu = 1$), and all other symbols have their usual meanings. The absorption rate of $\text{Ly}\alpha$ photons is proportional to $B_{jk} \int d\nu \phi(\nu) J_\nu$, where B_{jk} is the Einstein stimulated absorption coefficient from level j to level k , where we now mean to include the hyperfine structure of the atom. [The coefficients are related to the total spontaneous decay rates A_{kj} between the various hyperfine levels by $B_{jk} = (\lambda_{jk}^3 / 8\pi h) (g_k / g_j) A_{kj}$.] When the intensity varies smoothly across the H I absorption profile for $\text{Ly}\alpha$, that is, over at least one Doppler width $\Delta\nu_D = (b/c)\nu_\alpha$, where $b = (2kT_K/m_H)^{1/2}$ and ν_α is the line center frequency, it is useful to express the effect of the mixing mechanism in terms of the color temperature T_α of the radiation field, which we define by

$$\frac{1}{kT_\alpha} = -\frac{\partial \log \mathcal{N}_\nu}{\partial h\nu}, \quad (14)$$

evaluated at $\nu = \nu_\alpha$, where $\mathcal{N}(\nu)$ is the photon occupation number at frequency ν , $\mathcal{N}(\nu) = c^2 J_\nu / (2h\nu^3)$. If the background continuum spectrum near $\text{Ly}\alpha$ is a power law, $J_\nu \propto \nu^{-\alpha_S}$, then $T_\alpha = (h\nu_\alpha/k)/(3 + \alpha_S) \approx 3 \times 10^4 \text{ K} [4/(3 + \alpha_S)]$. It has been shown by Field (1959b), however, that, because of the large cross section for resonant scattering, the shape of the radiation spectrum near $\text{Ly}\alpha$ will be determined in all the relevant physical situations by the kinetic temperature of the scattering material. This tendency for thermalization is caused by the recoil of the atom, which shifts the profile to the red. As stimulated emission is negligible, the

radiation behaves like a classical relativistic Boltzmann gas, and the specific number density of photons, $n_\nu = 4\pi J_\nu / ch\nu$, assumes the quasi-LTE form $n_\nu = \text{const} \times \exp[-h(\nu - \nu_\alpha)/kT_K]$, where T_K is the kinetic temperature of the gas, to which the color temperature relaxes.

The influence of cosmological expansion on the Ly α profile has been discussed by Field (1959a). The effect of expansion is determined by two timescales, the mean-free scattering time for a Ly α photon, $b/(cn_{\text{HI}}\sigma_\alpha\lambda_\alpha)$, where n_{HI} is the density of neutral hydrogen, and the time it takes the Hubble expansion to redshift a photon through the resonance, $b/(Hc)$. Their ratio, also known within the framework of line radiation transfer theory as the Sobolev parameter (and equal in this context to the inverse of the Ly α optical depth), is

$$\gamma \equiv \frac{H}{\lambda_\alpha \sigma_\alpha n_{\text{HI}}} \approx 6 \times 10^{-6} h_{50} \left(\frac{1+z}{7} \right)^{-3/2} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^{-1}. \quad (15)$$

As this is much smaller than unity, the Ly α part of the radiation background spectrum establishes close thermal contact with the expanding IGM, and a quasi-static treatment of the cosmological line transfer is an adequate approximation (see, e.g., Rybicki & Dell’Antonio 1994). The profile of Ly α radiation in an expanding cloud has been calculated in detail by Deguchi & Watson (1985). These authors consider the case in which Ly α photons are created by recombinations inside the cloud. They show that $T_\alpha \rightarrow T_K$ for $\gamma \lesssim 10^{-5}$, similar to what is found in the absence of velocity gradients (Field 1959b). The color temperature approaches the gas kinetic temperature even more quickly if the cloud is irradiated by continuum radiation in the vicinity of Ly α , as this spectrum is already flat in frequency, and only a small number of scatterings is needed for the radiation field to achieve the proper slope (Deguchi & Watson 1985). In the next section we shall see how, in the presence of a strong ambient flux of Ly α photons, the color temperature governs the hyperfine populations of atomic hydrogen.

3.2. Ly α Photon Pumping

In the absence of collisions, the spin temperature, T_S , of the hyperfine levels of intergalactic hydrogen (as defined by the Boltzmann equation) is determined solely by the scattering of ambient

Ly α photons and by the microwave background radiation field at 21-cm (Field 1958, 1959a).² In a steady state we have

$$T_S = \frac{T_{\text{CBR}} + y_\alpha T_\alpha}{1 + y_\alpha}, \quad (16)$$

where $T_{\text{CBR}} = 2.73(1 + z)$ K is the temperature of the cosmic background radiation (Mather et al. 1994), and

$$y_\alpha \equiv \frac{P_{10}}{A_{10}} \frac{T_*}{T_\alpha} \quad (17)$$

is the normalized probability, or Ly α pumping “efficiency”. Here, $T_* \equiv h\nu_{10}/k = 0.07$ K, $A_{10} = 2.9 \times 10^{-15} \text{ s}^{-1}$ is the spontaneous decay rate of the hyperfine transition of atomic hydrogen, P_{10} is the indirect de-excitation rate of the triplet via absorption of a Ly α photon to the $n = 2$ level, and we have assumed $T_S \gg T_*$. In the absence of Ly α pumping the spin temperature goes to equilibrium with the 21-cm background radiation field on a timescale $T_*/(T_{\text{CBR}}A_{10}) \approx 5 \times 10^4 \text{ yr}$, and neutral intergalactic hydrogen cannot be seen in either 21-cm emission or absorption. If y_α is large, $T_S \rightarrow T_\alpha \approx T_K$, signifying equilibrium with the matter. The $1 \rightarrow 0$ transition rate via Ly α can be replaced by the total rate P_α ,

$$P_\alpha = \int d\Omega \int \frac{I_\nu}{h\nu} \sigma_\nu d\nu, \quad (18)$$

at which Ly α photons are scattered by an H atom in the gas, $P_{10} = 4P_\alpha/27$. For a point source, $I_\nu = f_\nu \delta(\Omega)$, where f_ν is the specific flux from the source. For an isotropic radiation field, $I_\nu = J_\nu$. In the limit $T_\alpha \gg T_{\text{CBR}}$, the fractional deviation in a steady state of the spin temperature from the temperature of the microwave background is

$$\frac{T_S - T_{\text{CBR}}}{T_S} \approx \left[1 + \frac{T_{\text{CBR}}}{y_\alpha T_\alpha} \right]^{-1}. \quad (19)$$

There exists then a critical value of P_α which, if much exceeded, would drive $T_S \rightarrow T_\alpha$. This “thermalization rate” is found to be

$$P_{\text{th}} \equiv \frac{27A_{10}T_{\text{CBR}}}{4T_*} \approx (5.3 \times 10^{-12} \text{ s}^{-1}) \left(\frac{1+z}{7} \right). \quad (20)$$

²This is not strictly true in the presence of bright continuum sources of 21-cm radiation, such as radio-loud quasars (see § 3.5.).

When $P_\alpha = P_{\text{th}}$, we have $T_S = 2T_{\text{CBR}}/(1 + T_{\text{CBR}}/T_\alpha)$. It should be emphasized that, while T_α depends on the spectral gradient across line center, $y_\alpha T_\alpha$, and hence P_{th} , are only functions of the magnitude of J_α , independent of the uncertainties regarding the actual profile of the Ly α background photon spectrum.

3.3. Level Mixing by Ly α Continuum Photons

There are four possible sources of background Ly α photons propagating into the (mostly) neutral material which has not yet been engulfed by an H II region: a) UV, non-ionizing continuum emitted by the radiating sources and then redshifted by the Hubble expansion to the Ly α wavelength in the absorbing medium rest-frame; b) Ly α photons generated by radiative recombinations and escaping from the H II regions; c) Ly α produced locally in the warm neutral material surrounding the ionized gas by radiative recombinations; and d) by collisional excitations. Of these, only the first, continuum photons redshifted into local Ly α , is a significant source of coupling, and we treat it next. We show that the remaining three are generally negligible in Appendix A.

In the limit where the scattering occurs close to the emission redshift, we find for a point source

$$P_\alpha \approx (2.3 \times 10^{-10} \text{ s}^{-1}) r_{\text{Mpc}}^{-2} L_{\alpha,47}, \quad (21)$$

where $L_{\alpha,47}$ is the source Ly α luminosity per octave in photon energy in units of 10^{47} ergs s^{-1} , and r_{Mpc} is the (physical) distance in Mpc between emission and scattering. Thus, within several Mpc from a bright, radio-quiet QSO, $P_\alpha > P_{\text{th}}$ and the spin temperature of neutral hydrogen will be determined by the Ly α continuum photons emitted from the quasar itself.

Alternatively, we may consider the effect of an integrated intensity from a distribution of sources. The Ly α scattering rate is then

$$P_\alpha \approx (8.5 \times 10^{-12} \text{ s}^{-1}) J_{\alpha,-21}, \quad (22)$$

where $J_{\alpha,-21}$ is the intensity of the background diffuse field at Ly α in units of

$10^{-21} \text{ ergs cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1}$. For $J_{\alpha,-21} \ll 1$, the spin system will go into equilibrium with the microwave background. Equation (22) is also approximately valid in the case of recombination $\text{Ly}\alpha$ photons, as radiative transfer calculations predict $dJ_\nu/d\nu = 0$ at the Doppler core of the transition (Urbaniak & Wolfe 1981). Spectral gradients induced by atomic recoil only result in higher order corrections for P_α . The rate is therefore dependent only on J_α .

The $\text{Ly}\alpha$ pumping mechanism poses a difficulty for detecting the IGM in absorption against the CBR, since, as we will show below, the same $\text{Ly}\alpha$ photons that mix the hyperfine levels of hydrogen heat the atoms as well. Before we discuss this problem, we briefly address the influence of particle collisions and direct radio continuum on the spin temperature.

3.4. Hyperfine Excitation via Particle Collisions

We demonstrate in this section that, except at very high redshifts, or in overdense regions, collisions are not effective in coupling the spin temperature to the kinetic temperature. One can show that, in the general case, the spin temperature may be written in the form (Field 1958)

$$T_S = \frac{T_{\text{CBR}} + y_\alpha T_\alpha + y_c T_K}{1 + y_\alpha + y_c}, \quad (23)$$

where

$$y_c \equiv \frac{C_{10}}{A_{10}} \frac{T_*}{T_K}, \quad (24)$$

and C_{10} is the rate of collisional de-excitation of the triplet level (cf. eqs. [16] and [17]). Emission occurs for

$$\frac{y_\alpha T_\alpha + y_c T_K}{y_\alpha + y_c} \approx T_K > T_{\text{CBR}}, \quad (25)$$

and absorption otherwise. Unless the electron density is larger than a few percent of the hydrogen density, the collisional de-excitation rate is dominated by collisions with other hydrogen atoms through spin-exchange (Purcell & Field 1956). Allison & Dalgarno (1969) give a maximum de-excitation rate of $C_{10}/n_{\text{H}} = 3.3 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$ for $T_K < 1000 \text{ K}$, with a value a factor 2.6 lower at $T_K = 100 \text{ K}$. To couple the spin temperature to the kinetic temperature requires a

collision rate $C_{10} > A_{10}T_{\text{CBR}}/T_*$, or a density $n_{\text{H}} > 3 \times 10^{-4}(1+z)\text{cm}^{-3}$. This may occur only for $1+z > 55(\Omega_{\text{IGM}}h_{50}^2/0.05)^{-1/2}$. At $z \sim 6$, collisions will mix the hyperfine levels if $\delta\rho/\rho \gtrsim 100$, i.e., in the case of nearly virialized halos.

3.5. Radio Loud Quasars

So far, we have limited our attention to radio quiet sources of radiation. However, 21-cm radiation emitted by quasi-stellar radio sources could raise the spin temperature of neighboring intergalactic H I above the CBR temperature (Bahcall & Ekers 1969; Urbaniak & Wolfe 1981). At a distance r from the QSO, the brightness temperature at 21-cm is

$$T_R = \frac{c^2}{8\pi k\nu_{10}^2} \frac{L_\nu}{4\pi r^2}, \quad (26)$$

where L_ν is the absolute luminosity of the radio-emitting QSO at 1.4 GHz. The typical radio quasar power is a small fraction of the B-band luminosity,

$$(\nu L_\nu)_{5\text{ GHz}} \approx 4 \times 10^{-4}(\nu L_\nu)_B \quad (27)$$

(e.g., Phinney 1985). So, assuming νL_ν is roughly constant in the radio, we derive

$$T_R \approx (30\text{ K}) \left[\frac{(\nu L_\nu)_B}{10^{47}\text{ ergs s}^{-1}} \right] r_{\text{Mpc}}^{-2}. \quad (28)$$

This temperature must be added to T_{CBR} in equation (16) for determining the level populations. Comparison with equation (21), however, shows that in the presence of UV photons this term is very small compared to the Ly α pumping term and so can safely be neglected.

4. Preheating of the IGM

While a cold IGM may be detected in absorption against the CBR, when the spin and kinetic temperatures are coupled through the Wouthuysen-Field mechanism, the period for which absorption is possible is short-lived. The reason is that the Ly α photons that couple the spin and kinetic temperatures will quickly heat the IGM through atomic recoil to a temperature above that

of the CBR. In this section, we return to the radiation sources introduced in § 2 and discuss their roles as heating sources by Ly α scattering and by photoelectric heating by soft X-rays.

4.1. Ly α Heating

The effect of recoil due to the finite mass of the atom was first discussed in resonance-line scattering by Field (1959b) (see also Adams 1971; Basko 1981). Contrary to the case of radiation scattered by free electrons (where a systematic shift in frequency due to recoil can produce large cumulative effects), in problems of resonance-line transfer a photon on the red side of the line has a heavy bias toward scattering back to the blue. This bias suppresses the additive effect of recoil: because of Doppler redistribution, a photon largely loses memory of its injection energy, and the background intensity develops only a slight asymmetry in frequency at line center (matching the slope of the Planck function in the Wien limit, see Field 1959b) even at the large Ly α optical depths normally encountered in astrophysical situations. However, it is through the recoil of the atom that the radiation field exchanges energy with the gas in scattering problems. In this section we shall discuss in some detail the effect of repeated Ly α resonant scattering on the thermal state of a cold, neutral IGM. A more formal derivation of the rate of transfer of energy from the radiation field to the atoms is provided in Appendix B.

The average relative change in a Ly α photon’s energy E after having been scattered by a hydrogen atom at rest is

$$\left\langle \frac{\Delta E}{E} \right\rangle = -\frac{h\nu_\alpha}{m_{\text{H}}c^2} \approx -10^{-8}, \quad (29)$$

where m_{H} is the mass of the hydrogen atom. It should be noted that this is an approximation valid only for $h\nu_\alpha \gg kT_K$. In the opposite limit, energy will flow from the atoms to the photons.

Through recoil, energy is transferred from photons to atoms at a rate

$$\dot{E}_\alpha = -\left\langle \frac{\Delta E}{E} \right\rangle h\nu_\alpha P_\alpha. \quad (30)$$

where P_α is the Ly α scattering rate per H atom. In the case of excitation at the thermalization

rate P_{th} , equation (30) becomes

$$\dot{E}_{\text{th}} = \frac{27}{4} \frac{(h\nu_{\alpha})^2}{m_{\text{H}}c^2} \frac{A_{10}T_{\text{CBR}}}{T_*} \approx (220 \text{ K Gyr}^{-1}) \left(\frac{1+z}{7} \right). \quad (31)$$

This rate is sufficient to drive the intergalactic gas kinetic temperature above the temperature of the cosmic background radiation in a fraction $\approx 15\%$ of the Hubble time at $z = 6$. As a consequence, a low density IGM is observable in absorption at 21-cm for only a brief interval of time once coupling is established. In general, coupling through the Wouthuysen-Field mechanism leads to emission.

This may be illustrated for the case of an early population of stars. It follows from equations (11) and (22) that the equivalent of only $\sim 10\%$ of the global present-day star formation rate is required for P_{α} to exceed P_{th} , and so drive $T_S \rightarrow T_K$. The IGM would initially be observed in absorption, but after $\sim 10^8$ yr, the IGM will be heated to a temperature above T_{CBR} by resonant scattering of the $\text{Ly}\alpha$ photons.

It is of interest to compare $\text{Ly}\alpha$ heating with heating by X-ray photons. The amount of energy deposited via photoionization by soft X-rays is, allowing for secondary ionizations, a few tens of eV per photon. There are also a few free electrons for every photon, as recombinations are negligible. In contrast, the total energy available for $\text{Ly}\alpha$ heating (whether the $\text{Ly}\alpha$ photon is part of the continuum or due to recombinations) is, neglecting the energy lost to cosmological expansion, given by the photon energy multiplied by the fractional line width. This is typically 0.001 eV per photon. Therefore, at least of order one $\text{Ly}\alpha$ photon per atom is needed to raise the kinetic temperature of the IGM above T_{CBR} . If we denote by L_X the soft X-ray luminosity of a young galaxy (associated with, e.g., supernova remnants, X-ray binaries, etc.), then photoelectric heating will dominate over $\text{Ly}\alpha$ heating if $L_X/L_{\alpha} \gtrsim 10^{-4}$. As the typical observed value for L^* galaxies is $L_X/L_{\alpha} \sim 10^{-2}$ (e.g., Fabbiano 1989; Bennet et al. 1994), $\text{Ly}\alpha$ photons from galaxies are an inefficient heating source compared to X-rays.³

³As shown by Rees (1985), the radiation pressure associated with trapped $\text{Ly}\alpha$ background photons is also unlikely to affect the condensing out of overdense regions of the IGM even in the linear regime.

4.2. Soft X-ray Heating

4.2.1. Heating Rate by Quasars

It is widely believed that the integrated ultraviolet flux arising from QSOs and/or hot, massive stars in metal-producing young galaxies is responsible for maintaining the intergalactic diffuse gas and the Ly α forest clouds in a highly ionized state at $z \lesssim 5$. In this and the next section we shall focus on a reionization scenario dominated by hard-spectrum quasar sources. In this case, while the radiation just shortward of the hydrogen Lyman edge will be absorbed at the ionization front generated by an individual QSO, very short wavelength photons will be able to propagate to much greater distances, $r_c > r_I$. The intensity of the QSO radiation flux at the light front will be reduced by the opacity of the (mostly) neutral gas surrounding the expanding H II region,

$$\tau_\nu = n_{\text{HI}}(r_c - r_I) \left(\sigma_\nu^{\text{HI}} + \chi \sigma_\nu^{\text{HeI}} \right), \quad (32)$$

where σ_ν^{HI} and σ_ν^{HeI} are the hydrogen and helium photoionization cross sections. For $\Omega_{\text{IGM}} h_{50}^2 = 0.05$, $z > 5$, and $r_c - r_I = 10 \text{ Mpc}$, the photoelectric optical depth will drop below unity for photon energies above 16 ryd. Thus, most of the photoelectric heating of the IGM will be accomplished by soft X-rays. The heating by such high energy photons, however, is not as efficient as for photons near the photoelectric threshold because of losses to collisional excitation of H I and He I and to the collisional production of secondary electrons (Spitzer & Scott 1969). The fraction f of photon energy converted into heat is then a sensitive function of the H II fraction. At the light front, this fraction is the residual ionization left over from the recombination epoch. For $h_{50} = 1$ and $\Omega_{\text{IGM}} = 0.05$, this is $\sim 5 \times 10^{-4}$ (Peebles 1993), which gives $f \approx 0.17$ (Shull & van Steenberg 1985). For a quasar with a spectral index α_S shortward of 1 ryd, the heating rate per hydrogen atom at the light front is given by

$$\dot{E}_X(r_c) = \alpha_S \frac{S}{4\pi r_c^2} h\nu_L f \int_1^\infty dx x^{-1-\alpha_S} [\sigma_\nu^{\text{HI}}(x-1) + \chi \sigma_\nu^{\text{HeI}}(x-1.8)] e^{-\tau_\nu}, \quad (33)$$

where $x \equiv \nu/\nu_L$, and ν_L is the frequency at the H I Lyman edge. Both the H I and He I photoelectric cross sections scale approximately as ν^{-3} far from threshold. In this

approximation the integral becomes exact, and may be expressed as

$$\dot{E}_X(r_c) \approx \left[(1400 \text{ K Gyr}^{-1}) \alpha_S \Gamma\left(\frac{2+\alpha_S}{3}\right) \tau_*^{(1-\alpha_S)/3} \right] \left(\frac{f}{0.17}\right) \left(\frac{S}{10^{57} \text{ s}^{-1}}\right) \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05}\right)^{-(2+\alpha_S)/3} \\ \times \left(\frac{1+z}{7}\right)^{-2-\alpha_S} \left(1 - \frac{r_I}{r_c}\right)^{-(2+\alpha_S)/3} \left(\frac{r_c}{25 \text{ Mpc}}\right)^{-(8+\alpha_S)/3}, \quad (34)$$

where $\Gamma(x)$ is the usual gamma function. Here, τ_* normalizes the opacity of the IGM for soft X-ray frequencies according to $\tau_\nu = \tau_*(\Omega_{\text{IGM}} h_{50}^2/0.05)[(1+z)/7]^3 (r_c/25 \text{ Mpc})(1-r_I/r_c)(\nu/\nu_L)^{-3}$. Numerically, $\tau_* \approx 7.8 \times 10^4$. For $\alpha_S = 1.5$, the coefficient in square brackets is $\sim 300 \text{ K Gyr}^{-1}$. The heating rate is dominated by helium absorption, which exceeds the hydrogen contribution by $\sim 2:1$. The intergalactic gas at the light front will be heated to a temperature above that of the cosmic background radiation in a fraction of the Hubble time,

$$\frac{\Delta t_{\text{heat}}}{t_H} \approx \left[0.02 \frac{\tau_*^{(\alpha_S-1)/3}}{\alpha_S \Gamma\{(2+\alpha_S)/3\}} \right] h_{50} \left(\frac{f}{0.17}\right)^{-1} \left(\frac{S}{10^{57} \text{ s}^{-1}}\right)^{-1} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05}\right)^{(2+\alpha_S)/3} \\ \times \left(\frac{1+z}{7}\right)^{9/2+\alpha_S} \left(1 - \frac{r_I}{r_c}\right)^{(2+\alpha_S)/3} \left(\frac{r_c}{25 \text{ Mpc}}\right)^{(8+\alpha_S)/3}. \quad (35)$$

For $\alpha_S = 1.5$, the coefficient in square brackets is ~ 0.1 . The H II region produced by a QSO will therefore be preceded by a warming front. Heating by soft X-rays is so efficient that the size of the warm zone will extend nearly to the light front. Note that, as the X-ray-heated bubbles around QSOs will survive as fossils even after the quasar has died, several generations N_g of quasars may actually be responsible for preheating the entire IGM. For a typical QSO age of $t_Q \approx 3 \times 10^8 N_g^{-1/3}$ yr, the required QSO comoving space density to heat the entire IGM to a temperature above that of the CBR by $z \approx 6$ is $\sim 10^{-10} \text{ Mpc}^{-3} N_g^{-1}$. By comparison, the comoving space density of bright QSOs at $z = 4$ is $\sim 100 N_g$ times larger (Warren, Hewett, & Osmer 1994). Note that, if all bright galaxies undergo a quasar phase, QSOs must have a very short lifetime, and $N_g \sim 100$. Soft X-rays from a few bright QSO sources could then prevent collapsing structures, such as protoclusters while still in the linear regime, from being detected in 21-cm absorption against the CBR (Scott & Rees 1990).

4.2.2. Numerical Solution for Expanding Warm H II Regions Around Quasars

We illustrate the thermal effect of an expanding H II region on the surrounding medium by numerically solving the time-dependent equations for photoionization and heating around a point source which turns on suddenly in a homogeneous, expanding universe. We have integrated as a function of time the ionization equations for hydrogen and helium, including photoionization, collisional ionization, and radiative recombination, on a radial grid with a local ionization flux equal to $L_\nu \exp(-\tau_\nu)/4\pi r^2$, where τ_ν is given by equation (32). The flux is set to zero beyond the light front to prevent unphysical excess heating at large distances. We use the time-dependent ionization algorithm of Meiksin (1994), and include here the generation of secondary electrons by soft X-rays. We adopt the fitting formulas of Shull & van Steenberg (1985) to include the additional ionization and energy losses associated with the production of these electrons. We solve the time-dependent energy equation along with the ionization equations for a gas of primordial hydrogen and helium composition, including photoelectric heating and radiative losses due to collisional excitation and ionization, radiative recombination, Compton cooling off the CBR, and cooling by expansion. An accurate numerical solution for the growth of the H II region requires grid zones that are optically thin to the ionizing radiation when the hydrogen is still fully neutral. Thicker zones are ionized too rapidly, resulting in a too rapidly growing I-front. Because we wish to follow the heating out to the light radius, however, this restriction imposes a prohibitive demand on our computing resources. The demand imposed by helium ionization is less restrictive because of its lower density, and so may be met. We therefore adopt a compromise. A grid sufficiently fine to resolve the helium ionization is used, but it is somewhat too coarse to follow the evolution of the H II I-front accurately. To prevent too rapid a growth of the H II front, we set the flux to zero for energies less than 1.8 ryd, the He I ionization potential. We find this results in accurate placement of the H II front, though the details of its structure may not be quite correct. Because helium ionization and the soft X-ray heating are governed by the opacity outside the H II front, this approach yields an accurate solution for the He II and He III fronts and the external heating rate in the surrounding neutral IGM.

In Figure 1, we show the results of an integration for a QSO turning on at $z = 6$ with a gradual rise over a time of 10^7 yr to a peak photoionizing luminosity of $S = 10^{57} \text{ s}^{-1}$, and with $\alpha_S = 1.5$, in a medium with $\Omega_{\text{IGM}} h_{50}^2 = 0.05$. By $z = 5.5$, the QSO has heated the surrounding gas to a temperature exceeding that of the CBR to a (proper) distance of 23 Mpc, compared to the light front radius of 25 Mpc. The temperature climbs inward, reaching 1000 K at a distance of 10 Mpc from the QSO. The ionization fronts are at 7.5 Mpc at this time. Neither the He II nor the He III fronts extend much beyond that of hydrogen, but for different reasons. Because of the much higher photon energy required to photoionize He II compared to H I, the radius of the He III front can exceed that of H II by at most a factor of $(4^{-\alpha_S}/\chi)^{1/3} \simeq 1.1$, despite the low helium abundance (cf. Madau & Meiksin 1994). While the energy needed to photoionize He I is smaller, the IGM is more optically thick at the lower energy. At $z = 5.5$, the opacity of H I at the He I edge reaches unity a distance 10 kpc beyond the H II ionization front. Outside the H II region, H I efficiently shields the helium from the ionizing radiation above 1.8 ryd, ensuring that it remains mostly neutral. The He I is then able to warm through the absorption of the incident soft X-rays emitted by the QSO.

4.2.3. Heating Rate by Virialized Halos

Galactic mass halos provide another source of heating in CDM-dominated cosmologies. While before stars form the numbers of these halos are small, too small to account for the ionization of the IGM by UV photons, we demonstrate here that a sufficient number of halos will collapse to provide the soft X-rays required to heat the IGM to a temperature above that of the cosmic background radiation.

We start with equation (9), which gives the background emissivity above a minimum photon energy. To specify its value, we observe that, at sufficiently low energies, the distance a photon travels before being absorbed in the IGM will be shorter than the mean separation of the sources, so that only part of the IGM will be penetrated by such photons. We set the minimum energy according to the criterion that the filling factor of photons with energy above $h\nu_{\text{min}}$ be greater

than unity. For $5 < z < 10$, this yields

$$h\nu_{\min} \approx 15.2 \text{ ryd} \left(\frac{1+z}{7} \right)^{3/2}, \quad (36)$$

corresponding to a halo of comoving size $\approx 0.9 \text{ Mpc}$, virial temperature $\approx 1.5 \times 10^6 \text{ K}$, and total mass $\approx 2 \times 10^{11} \text{ M}_\odot$. Approximated in this manner, the emissivity is only half of the value given in equation (9), since halos with $r_0 < \hat{r}_0$ are excluded by the filling factor criterion. Note that the collisional ionization level of helium in such hot halos is so large that absorption of soft X-rays within the halos themselves is completely negligible. Substituting equation (36) into equation (9), and taking into account the factor of 2, we find for the heating rate of the halos per hydrogen atom

$$\dot{E}_X^{\text{halo}} \approx (2.6 \times 10^5 \text{ K Gyr}^{-1}) \left(\frac{f}{0.2} \right) \left(\frac{1+z}{7} \right)^3 \exp \left[-6.6 \left(\frac{1+z}{7} \right)^{1.6} \right], \quad (37)$$

where $f \approx 0.2$ accounts for the energy loss to collisional excitations and ionizations per incident X-ray photon (Shull & van Steenberg 1985), for an assumed H II fraction of 10^{-3} . We show the heating rate in Figure 2. The IGM will be heated to a temperature above that of the CBR in a fraction of the Hubble time,

$$\frac{\Delta t_{\text{heat}}^{\text{halo}}}{t_H} \approx (7.5 \times 10^{-5}) \left(\frac{f}{0.2} \right)^{-1} \left(\frac{1+z}{7} \right)^{-0.5} \exp \left[6.6 \left(\frac{1+z}{7} \right)^{1.6} \right]. \quad (38)$$

At $z = 6$, this fraction is 0.05. Because the merging of small halos into larger ones is an ongoing process, the full Hubble time will not be available to the halos at any given stage in the hierarchy for heating the IGM. Moreover, the halos will cool by radiative processes on a timescale of $t_{\text{cool}}/t_H \simeq 0.7[(1+z)/7]^{-3/2} T_6^{1.8}$ (Gould & Thakur 1970), and even more quickly at high redshifts by Compton cooling off the CBR. The cooling, however, may not be significant for $z < 10$ before the gas is reheated by mergers in the next stage of the hierarchy. Allowing the halos to provide X-ray photons for 50% of the Hubble time, they will heat the IGM to a temperature of $\sim 100 \text{ K}$ by $z = 6$. At higher redshifts, halo heating may still be important in overdense environments like protoclusters where the number density of halos may be substantially boosted by statistical bias (Bardeen et al. 1986). Of course once stars start to form, they or, more likely, their remnants may dominate the soft X-ray emission of an individual halo. In this sense, the halo emissivity calculations outlined above present a lower limit to the heating rate of the IGM at early epochs.

It is worth noting that while the X-rays are an efficient heating mechanism of the gas, they are an inefficient source of photoionization. This is because the time required to photoionize the IGM by X-ray photons is larger by a factor of $\sim (15 \text{ ryd}/kT_{\text{CBR}}) \approx 10^5$, since now it is the number of ionizing photons rather than their energy that counts. The ionization time thus exceeds the Hubble time by a few orders of magnitude, and the IGM will remain mostly neutral.

4.2.4. Thermal Coupling Between Neutral Atoms and Ions

In the discussion above, we have assumed that the energy deposited by the photons is strongly coupled to the hydrogen. The thermal energy, however, is initially contained in the photoejected electrons and ions only. In this section we shall verify that the equilibration mechanisms between the various ions and neutrals in the low density IGM are sufficiently rapid to ensure that all species are thermally coupled.

We first consider Coulomb coupling between the ions. The time for two different species i and j to reach thermal equilibrium through Coulomb scattering is

$$t_{\text{eq}}(i, j) \approx (5.87 \text{ s}) \frac{A_i A_j}{(n_i + n_j) Z_i^2 Z_j^2 \log \Lambda} \left(\frac{T_i}{A_i} + \frac{T_j}{A_j} \right)^{3/2}, \quad (39)$$

(Spitzer 1962). Each species has a temperature T , density n , charge Z , and molecular weight A ($A = 1/1836$ for electrons). (This expression applies strictly only when the total energy between the two species, $n_i T_i + n_j T_j$, is conserved during the equilibration process. For a single species, $j = i$ may be set in equation (39) if the coefficient 5.87 is replaced by 8.06.) In the physical regime of interest, $20 < \log \Lambda < 50$. Once the electrons and protons each reach equilibrium separately (the shortest two timescales), the bottleneck for all the ions to relax to a single temperature is proton-electron scattering, with a coupling time

$$t_{\text{eq}}(e, p) \approx (3.3 \text{ Myr}) \left(\frac{x_e}{10^{-3}} \right)^{-1} \left(\frac{1+z}{7} \right)^{-3} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^{-1} T_4^{3/2}, \quad (40)$$

where x_e is the number of electrons per hydrogen nucleus. After these two species equilibrate, singly and doubly ionized helium follow shortly through proton scatters. Note that, as the

post-photoionization (by 16 ryd soft X-rays) temperature of electrons is $\sim 10^6$ K, the time for protons to thermalize with such high energy electrons by Coulomb collisions alone actually exceeds the Hubble time. At these energies, however, the primary photoelectrons will quickly cool via the production of secondary electrons by the collisional ionization and excitation of neutral hydrogen (Spitzer & Scott 1969). The cooling rates of an electron of energy E ryd by these two processes are comparable, and are given by

$$t_{\text{cool}} \approx (0.03 \text{ Myr}) \left(\frac{1+z}{7} \right)^{-3} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^{-1} \frac{E^{3/2}}{\log E}, \quad (41)$$

(Shull & van Steenberg 1985). For $E = 16$, this gives $t_{\text{cool}} \approx 0.6$ Myr. The result is an energy distribution of primary and secondary electrons which declines sharply above the $\text{Ly}\alpha$ excitation energy of 10.2 eV (Bergeron & Souffrin 1971).

The coupling rate is still small for electrons of these energies, but they further lose energy through elastic scattering off neutral hydrogen atoms. The electron-hydrogen cross section for a 10 eV electron is about $7\pi a_0^2$, where a_0 is a Bohr radius (Moiseiwitsch 1962). The corresponding relaxation timescale is

$$t_{\text{eq}}(e, \text{H}) \approx (25 \text{ Myr}) \left(\frac{1+z}{7} \right)^{-3} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^{-1}. \quad (42)$$

The electrons can finally cool to temperatures below $\sim 10^4$ K. Once a thermal distribution of electrons has been established, new energetic electrons generated by X-ray photoionization will reach equipartition with the cooler electrons on a much shorter timescale by Coulomb losses (Shull & van Steenberg 1985).

The distance from the light-front for which the temperature of the neutral hydrogen atoms lags behind the electron temperature corresponding to equation (42) is ~ 8 Mpc. This, however, assumes complete thermalization. The less stringent requirement that the neutral hydrogen atoms be heated to a temperature large compared to the cosmic background radiation temperature will be satisfied on a somewhat shorter timescale. The number of QSOs required to heat the IGM above the cosmic background radiation temperature may then exceed the estimate given in § 4.2.1 by a factor of a few only.

The neutrals will reach a single common temperature relatively quickly. The interaction between neutral hydrogen atoms is dominated by elastic scattering.⁴ The equilibration time is $t_{\text{eq}} = 2/n_{\text{HI}}\langle u\sigma \rangle$, where $\langle u\sigma \rangle$ is the slowing down coefficient for hydrogen atom collisions (Spitzer 1978). Using the elastic collision cross section from Allison & Smith (1971), this gives

$$t_{\text{eq}}(\text{H}, \text{H}) \approx (0.4 \text{ Myr}) \left(\frac{1+z}{7} \right)^{-3} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^{-1} T_4^{-1/4}. \quad (43)$$

Thus all the species will reach a common temperature within a timescale much shorter than the Hubble time, and it is safe to assume that the hydrogen atoms may be described by a single Maxwellian energy distribution.

5. Observations of Redshifted 21-cm Line Radiation

In this section we discuss how radio measurements may reveal 21-cm line radiation from a neutral IGM at high redshift. The *Giant Metrewave Radio Telescope* (GMRT), scheduled for completion by 1997, has two frequency bands which are best suitable for detecting 21-cm radiation at $z > 5$, viz. 150 ± 20 and 235 ± 8 MHz (Swarup 1996). The central frequencies of these bands correspond to $z = 8.5$ and 5.0 , respectively. Because of its high sensitivity for both compact and extended sources, the GMRT will be a valuable instrument for cosmological studies in the redshift ranges $4.85 < z < 5.25$ and $7.35 < z < 9.9$. We focus the discussion on measurements in the 150 MHz band, corresponding to $z \approx 8$.

We have assumed that the IGM at early epochs is uniform on large scales. One general issue is the effect of small-scale clumping. In a CDM-dominated cosmology, inhomogeneity on subgalactic scales develops before the first quasars form (if the latter require galactic-mass collapsed structures). This may delay the eventual percolation of H II regions, and increases the value of J needed to balance recombinations and maintain full ionization. It is unlikely that

⁴Provided that statistical equilibrium has been established between the singlet and triplet ground state levels. This is achieved on a timescale $\sim 5 \times 10^4$ yr by collisions with CBR photons.

clumping would change the present discussion by very much. It leaves unchanged the amount of heating and ionization due to each UV photon, as well as the rate at which the heating and ionization fronts advance around a quasar. Even in an overdense region (unless the overdensity is extreme), recombinations are unimportant over the lifetime of a single quasar (or during the time it takes for the heating front to advance). It is also worth emphasizing that there will unavoidably be nonuniformities in the 21-cm intensity due to density inhomogeneities, but that these are restricted to smaller scales than those due to the patchy heat input (at least in the QSO case). While we confine our discussion to larger scales, it would nonetheless be worthwhile to consider means for their detection.

To illustrate the basic principle of the observations we propose, consider a region of neutral material with spin temperature $T_S \neq T_{\text{CBR}}$, having angular size on the sky which is large compared to a beamwidth, and radial velocity extent due to the Hubble expansion which is larger than the bandwidth. Its intergalactic optical depth at $21(1+z)$ cm along the line of sight,

$$\tau(z) = \frac{3c^3 h^3 n_{\text{HI}}(0) A_{10}}{32\pi H_0 k^3 T_*^2 T_S} (1+z)^{1.5} \approx 10^{-2.9} h_{50}^{-1} \left(\frac{T_{\text{CBR}}}{T_S} \right) \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right) (1+z)^{1/2}, \quad (44)$$

will typically be much less than unity. The experiment envisaged consists of two measurements, separated in either angle or frequency, such that one measurement, the fiducial, detects no line feature, either because there is no H I or because $T_S \approx T_{\text{CBR}}$, and the second at $T_S \neq T_{\text{CBR}}$. As the brightness temperature through the IGM is $T_b = T_{\text{CBR}} e^{-\tau} + T_S(1 - e^{-\tau})$, the differential antenna temperature (which is the relevant quantity since we are considering a detector operating as a comparison system), observed at Earth between this region and the CBR will be

$$\delta T = (1+z)^{-1} (T_S - T_{\text{CBR}}) (1 - e^{-\tau}) \approx (0.011 \text{ K}) h_{50}^{-1} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right) \left(\frac{1+z}{9} \right)^{1/2} \eta, \quad (45)$$

where we have defined a “21-cm radiation efficiency” as

$$\eta \equiv x_{\text{HI}} \left(\frac{T_S - T_{\text{CBR}}}{T_S} \right). \quad (46)$$

Here x_{HI} refers to the neutral fraction of the hydrogen in the region for which $T_S \neq T_{\text{CBR}}$. As long as T_S is much larger than T_{CBR} (hence if there has been significant preheating of the intergalactic

gas), $\eta \rightarrow x_{\text{HI}}$, and the IGM can be observed in emission at a level which is independent of the exact value of T_S . By contrast, when $T_{\text{CBR}} \gg T_S$ (negligible preheating), the differential antenna temperature appears, in absorption, a factor $\sim T_{\text{CBR}}/T_S$ larger than in emission, and it becomes relatively easier to detect intergalactic neutral hydrogen (Scott & Rees 1990). Note that in a universe with $q_0 \ll 1/2$, δT increases more nearly linearly with $(1+z)$, and the numerical coefficient in equation (45) may be larger by up to a factor $3[(1+z)/9]^{1/2}$. Depending on the cosmology, δT will increase toward low frequencies at a rate between $\nu^{-1/2}$ and ν^{-1} .

The discrete nature of the sources of preheating and 21-cm excitation will cause δT to have a distinctive structure both in frequency and in angle across the sky, as will be shown below. It is important to point out that the radio signal from the IGM will be swamped by the much stronger non-thermal background that dominates the radio sky at meter wavelengths. This synchrotron-type component has brightness temperature

$$T_b(z=0) \approx 6 \text{ K} \left(\frac{\lambda}{1 \text{ m}} \right)^{2.8}, \quad (47)$$

which is consistent with the integrated contribution from discrete radio sources (see Longair 1995 for a recent review). The anisotropic galactic contribution is about three times larger. Such emission will cause intensity fluctuations from beam to beam, either from discrete radio sources or from variable galactic brightness. Hence the difficulty is to be able to discriminate between the IGM around a QSO with a brightness temperature of 0.01 K against a small fluctuation in the galactic and extragalactic background. Clearly, to remove this foreground contamination, some differential mapping technique must be used. The line radiation from high- z gas might, e.g., be distinguished against continuum emitting sources owing to its sharp and distinctive spectral feature, i.e., by making measurements at two frequencies sufficiently close that the continuum emission from discrete sources does not vary significantly (Bebbington 1986).

In the following sections we will illustrate by a few examples the sensitivity of the detailed pattern of sky brightness at redshifted 21-cm wavelengths to the unknown thermal and ionization history of the universe.

5.1. Quasars in a Background of UV Radiation at $z \approx 8$

In the theories for the origin of cosmic structure that are most popular today, massive objects grow by hierarchical clustering – low mass perturbations collapse early, then merge into progressively larger systems. In such a scenario, one naturally expects the high redshift universe to contain small scale substructures, systems with virial temperature well below 10^4 K. The resultant virialized systems will remain as neutral gas clouds unless they can cool due to molecular hydrogen, in which case they may form Pop III stars. These stars will then ionize the gas in their vicinity. It is less clear, however, whether or not a sufficient number of ionizing photons will be able to penetrate all of the surrounding neutral gas in the cloud to reach the external intergalactic gas and ionize it. Nonetheless, the stars will still produce a background of UV continuum photons near the $\text{Ly}\alpha$ frequency which can escape.

It is therefore of interest to consider a case in which QSOs turn on in a universe where the $\text{Ly}\alpha$ scattering rate is $P_\alpha \approx P_{\text{th}}$ everywhere, i.e., where

$$J_\alpha = J_{\text{th}} \equiv \frac{9}{2} \frac{A_{10}\nu_\alpha^3}{A_\alpha\nu_{10}} \frac{kT_{\text{CBR}}}{c^2} \approx (10^{-21.1} \text{ ergs cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1}) \left(\frac{1+z}{9} \right), \quad (48)$$

(cf. eqs. [20] and [22]). The $\text{Ly}\alpha$ photons will propagate into uncollapsed regions of the IGM where the kinetic temperature in the absence of heating will be

$$T_K(z) \approx 0.026 \text{ K}(1+z)^2 \quad (49)$$

(Couchman 1985), well below T_{CBR} because of adiabatic cooling during cosmic expansion. The $\text{Ly}\alpha$ flux will couple the spin temperature to the kinetic temperature, and T_S will be pulled below T_{CBR} . (Note that, in the absence of a radiative coupling agent like $\text{Ly}\alpha$, this could only happen in very high-baryon-density models, where collisions would remain important even at modest redshifts, Scott & Rees 1990.)

As shown in § 4.1, however, the same $\text{Ly}\alpha$ photons which mix the hyperfine levels will also heat the gas through recoil. Integrating the energy equation, assuming the $\text{Ly}\alpha$ field is turned on, or more precisely reaches the level of equation (48), at redshift z_α , and using equations (31) and

(49), we find

$$T_K(z) \approx (0.026 \text{ K})(1+z)^2 + (160 \text{ K})(1+z)^2[(1+z)^{-5/2} - (1+z_\alpha)^{-5/2}]. \quad (50)$$

The characteristic timescale for heating the medium above the CBR temperature via Ly α resonant scattering at the thermalization rate is

$$\Delta t_{\text{heat}} = \frac{2}{9} \frac{m_{\text{H}} c^2 \nu_{10}}{h \nu_\alpha^2} A_{10}^{-1} \approx 10^8 \text{ yr}, \quad (51)$$

about 20% of the Hubble time at $z \approx 8$. The heating rate corresponding to equation (48) is shown in Figure 2. The result is a finite interval of time during which Ly α photons couple the spin temperature to the kinetic temperature of the IGM before heating the IGM above the CBR temperature. If Ly α sources turned on at redshifts $z_\alpha \lesssim 10$, this interval would present a window of opportunity in redshift space near $z \approx 8$ that would enable a large fraction of intergalactic gas to be observable at ~ 160 MHz in *absorption* against the CBR.

In such a scenario, 21-cm emission on tens of megaparsecs scale would still be produced by the relatively rare, X-ray heated regions associated with QSO sources. If we denote by $T_{S,w}$ the spin temperature of the “warm” ($T_{S,w} \gg T_{\text{CBR}}$) gas, and by $T_{S,c}$ the spin temperature of the surrounding, more pervasive “cold” ($T_{S,c} < T_{\text{CBR}}$) medium, as given by equation (50), the differential antenna temperature observed between these two phases of the IGM will be proportional to the effective efficiency

$$\Delta\eta = T_{\text{CBR}} \left(\frac{x_{\text{HI},c}}{T_{S,c}} - \frac{x_{\text{HI},w}}{T_{S,w}} \right). \quad (52)$$

Here, we distinguish between the neutral fractions in the cold and warm regions, to allow the warm zone to refer either to the inside or the outside of the ionized bubble. Even though the H I in the cold region is absorbing the CBR, by using the flux through the cold gas as a fiducial, the signature of the cold gas will appear as *emission* from the warm regions surrounding the quasars. The effective efficiency will remain much larger than unity, $\Delta\eta \approx T_{\text{CBR}}/T_{S,c} \gtrsim 4$, only for $\lesssim 10^7$ yr (cf. equation [50]), to drop to $\Delta\eta \approx 2$ after $\approx 5 \times 10^7$ yr. Note that $\Delta\eta \approx 4$ corresponds to temperature fluctuations of about 0.04 K. This flux is well below the level of the Bebbington’s

(1986) radio survey, but would be reachable with an instrument like the GMRT in several hours of integration.

For illustrative purposes, consider now the sphere of influence associated with a quasar ($L_{\alpha,47} = 1$, $\alpha_S = 1.5$) which turned on at $z = 8.7$, nearly synchronized with a Ly α background of sources turned on at $z = 9$, and radiating isotropically. At $z = 8$, e.g., after 5×10^7 yr the QSO will have ionized a bubble of size $r_I \approx 7.5$ Mpc, and heated the surrounding H I within $r_c \approx 15$ Mpc. In Figure 3a we show $\Delta\eta$ as a function of radius at various epochs, including the effect of Ly α heating on the cold phase. The peak value at the I-front initially exceeds unity, but declines as $T_{S,c}$ rises. The quasar has created a shell of 21-cm emission surrounding the H II region with a thickness of $\Delta r \approx 7.5$ Mpc. The hydrogen column density through the shell's width is $\approx 2 \times 10^{21} \text{ cm}^{-2}$, and its total H I mass is $\approx 2 \times 10^{16} M_\odot$. The angular size on the sky corresponding to the diameter of the emitting region is

$$2\theta = \frac{H_0(1+z)^2 r_c}{c(1+z - \sqrt{1+z})} \approx 2 \text{ degrees.} \quad (53)$$

The shell will have a characteristic radial double peaked profile, as the gas in its interior will be completely ionized, as shown in Figure 3b. The velocity difference across each peak, due to the Hubble flow, is $\Delta v = H_0(1+z)^{3/2} \Delta r \approx 10,000 \text{ km s}^{-1}$. Between the inner edges of the two peaks, Δv is twice as large, since $\Delta r \approx r_I$ at this time. At 150 MHz, we would then observe two signals (the front and back of the shell) with flux density

$$\delta S \approx (6 \text{ mJy}) h_{50}^{-1} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right) \left(\frac{1+z}{9} \right)^{-3/2} \left(\frac{\Delta\eta}{2.2} \right) \quad (54)$$

(cf. eqs. [45] and [52]), extending over and separated by about 5 MHz, nearly the entire bandwidth of the receiver. For a fixed velocity width, δS increases linearly with the mass of the shell.

Let us analyze in some detail the detectability of a typical 21-cm emitting shell with the GMRT. The 30 antennas of the GMRT will be configured as a Y of 16 dishes with a central compact core of 14 dishes. Because the GMRT functions as an interferometer, it is insensitive to angular scales exceeding $\sim \lambda/d$, where d is the separation of the dishes. If we wish to measure ~ 2 degree-scale structures at 150 MHz ($\lambda \sim 2 \text{ m}$), the projected separation of antennas is required

to be about $d \lesssim 50$ m. Since the GMRT antennas have a diameter of about 45 m, the GMRT interferometric system will not be sensitive to 2 degree-scale structures as the antennas will start shadowing each other for projected spacings less than 45 m.

An alternative strategy is to measure the total power by incoherent addition of the outputs of the 30 GMRT antennas, for which the effective collecting area will be 30×0.5 that of one antenna. For observations at 150 MHz the theoretical rms thermal noise of the GMRT output in incoherent mode is ≈ 1 mJy for the following parameters: $T(\text{sys}) \sim 600$ K, $A(\text{eff}) \sim 1000 \text{ m}^2$ for 1 45-m dish, a bandwidth of $10,000 \text{ km s}^{-1}$ (5 MHz), and 5-hr integration for each of the 2 frequency bands separated by about 5 MHz (Swarup & Malik 1996). A comparison with equation (54) shows that ~ 100 hours of integration will be required to detect an individual shell at $5\sigma_{\text{rms}}$. This time is comparable to that proposed for searches for emission from collapsing H I clouds at high redshift (Swarup et al. 1991; Subramanian & Padmanabhan 1993; Swarup 1996). However, it is not clear whether the above sensitivity can be realized in practice because of the man-made noise and variations of the instrumental gain and the galactic background temperature over 2 bands separated by 5 MHz. Note that, although the differential antenna temperature will be larger at earlier epochs because of the larger brightness contrast (i.e., less time to raise the temperature of the cold phase), the smaller radius of the QSO sphere of influence more than cancels the increase in emission efficiency. The signal-to-noise ratio improves slightly with increasing bubble size.

To calculate the expected number of emitting shells in a radio survey, we may estimate their effective covering factor on the sky as

$$C = \frac{c}{H} n_{\text{QSO}} d_A^2 (\theta + \Theta)^2 \frac{\Delta v}{c}, \quad (55)$$

where n_{QSO} is the (proper) number density of quasars, d_A is the angular-diameter distance, Θ is the field of view of the radio antenna (about 3.1 degrees at 150 MHz), and Δv measures the radial dimension covered by the survey. If the latter matches the shell thickness Δr , e.g., for a bandwidth of 5 MHz, the covering factor may be rewritten as

$$C = q_{21} \left(1 + \frac{\Theta}{\theta}\right)^2 \frac{\Delta r}{r_c}, \quad (56)$$

where q_{21} is the volume filling factor of the 21-cm emitting bubbles. As $\Theta^2 \Delta r / \theta^2 r_c \approx 5$, it follows from equation (56) that a shell will be detected after scanning about 20 fields with the GMRT if the filling factor of warm bubbles is $\approx 1\%$.

5.2. Isolated Quasars

We shall now analyze the model problem of an individual, isolated QSO turning on at epoch $z > 5$ in an otherwise unperturbed IGM. In this case the only radiative coupling agent is the continuum UV radiation of the quasar itself, redshifted to $\text{Ly}\alpha$. From equations (21) and (20), it is useful to define a “thermalization distance” from a QSO, denoted by r_{th} , at which $P_\alpha = P_{\text{th}}$. Numerically, this yields

$$r_{\text{th}} \approx (6 \text{ Mpc}) \left(\frac{1+z}{9} \right)^{-1/2} L_{\alpha,47}^{1/2}. \quad (57)$$

For the best case of $T_\alpha \gg T_{\text{CBR}}$, the efficiency for 21-cm emission may be expressed, using equations (17), (19), (20), and (46), as $\eta = (1 + r^2/r_{\text{th}}^2)^{-1}$. At $r = r_{\text{th}}$, $\eta = 1/2$. Comparison with equation (3) shows that the thermalization distance is comparable to the size of the H II region. Since $r > r_I$ is required for detection, $\eta > 1/2$ may be achieved only for a brief redshift interval after the quasar turns on. This interval is so small for many quasars that their H II regions will still be in the luminal expansion stage (cf. equation [2]), so that they will not be preceded by a warming front and will thus not be visible in 21-cm emission. The efficiency will therefore generally lie in the range $0 < \eta < 1/2$.

Larger shells offer a slight advantage for detectability. This is because the flux of a shell emitting 21-cm radiation depends not only on its efficiency but on its size as well. The frequency integrated flux is proportional to

$$(\Omega \Delta \nu)_{\text{eff}} \equiv 2\pi \int d\nu \int d\theta \theta \eta. \quad (58)$$

Consider a measurement made at a frequency corresponding to a depth in the emitting zone around the quasar just outside the I-front, and with a bandwidth $\Delta \nu_c$ matched to the light radius r_c . Denoting by θ_{th} the angle on the sky subtended by r_{th} , integrating equation (58) out to r_c ,

using the expression above for η , gives $(\Omega\Delta\nu)_{\text{eff}} \simeq 2\pi\theta_{\text{th}}^2\Delta\nu_c[1 - (r_{\text{th}}/r_c)\arctan(r_c/r_{\text{th}})]$, when $r_c \gg r_I$. While the signal increases initially like the cube of the size of the emitting region, it does so only linearly at later times. Since the noise increases like $r_c^{3/2}$, the signal-to-noise ratio increases asymptotically only like $r_c^{1/2}$. Allowing for a light radius that subtends an angle θ_c on the sky of 1° , the flux is then

$$\delta S \approx (1 - 2 \text{ mJy}) h_{50} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right) \left(\frac{1+z}{9} \right)^{-1/2} L_{\alpha,47} \quad (59)$$

over a bandwidth of $\Delta\nu_c \approx 11 \text{ MHz} [(1+z)/9]^{-1/2}(\theta_c/1^\circ)$, comparable to the rms noise. It would thus take a few hundred hours of integration time on the GMRT to detect the shell at $5\sigma_{\text{rms}}$. A quicker strategy may be to search for an excess variance across the sky over the rms noise that these shells would produce in the observed flux per resolution (beamwidth \times bandwidth) element. A next generation facility like the *Square Kilometer Array Interferometer* (Braun 1996) could easily map warm IGM bubbles, effectively opening much of the universe to a direct study of the reheating epoch.

6. Summary

The spectral appearance at 21-cm of non-uniform gas at high- z had been previously investigated by Hogan & Rees (1979) and Scott & Rees (1990). These earlier papers showed that the 21-cm emission from high- z H I may display angular structure, as well as structure in redshift space; radio observations could therefore in principle yield “tomographic” mapping of protoclusters or other large-scale inhomogeneities at high redshifts, where the gas had not yet been reionized. The present work discusses how similar techniques may be used to reveal “patchiness” in the 21-cm emission due to non-uniformities in the spin temperature, rather than in the density. Such effects could tell us how (as well as when) the primordial gas was reheated; if the reheating were due to ultraluminous but sparsely distributed sources (e.g., the first QSOs), the resultant patches would have a larger scale than any gravitationally-induced inhomogeneities, and would therefore be less difficult to detect. The continuum radiation, redshifted to Ly α , from an early

generation of stars would couple the spin temperature to the low temperature of the IGM for a transitory interval before heating the gas above the CBR temperature. The result is a “window of opportunity” during which it may be possible to detect the IGM in absorption against the CBR, and in so doing possibly reveal the first epoch of star formation.

We have assessed three preheating mechanisms at early times: soft X-ray radiation from quasar sources, soft X-rays from virialized $10^{10} - 10^{11} M_{\odot}$ halos, and resonant scattering of background $\text{Ly}\alpha$ continuum photons from an early generation of stars, and argued that radiation sources at high redshifts may prevent collapsing structures that cannot cool efficiently from being detected in 21-cm absorption against the microwave background. We have shown that scattering of $\text{Ly}\alpha$ continuum radiation can mix the hyperfine levels of intergalactic H I within ~ 10 Mpc of an individual QSO, and that this will produce an appreciable 21-cm emissivity from the IGM in the quasar neighborhood. While the X-ray-heated bubbles will still survive as fossils, once the quasar has died, the $\text{Ly}\alpha$ photons will rapidly adiabatically redshift out of resonance so that no emission will occur without an external supply of photons.

If reionization occurred at redshift $5 \lesssim z \lesssim 10$, these QSO spheres of influence might be detectable at meter wavelengths by the GMRT. Because of the very large scales involved, there is no difficulty in achieving a sufficiently narrow bandwidth and an adequate angular resolution with the GMRT; the main limitations are the low flux levels and the radio background subtraction.

It is possible that both the reionization and heating of the IGM were due to an early population of stars on subgalactic scales (rather than quasars) at redshift $z \gtrsim 20$, well beyond the reach of the GMRT. Even if they reionized the IGM at an epoch accessible to the GMRT, however, these sources would be less thinly spread, and the characteristic structures in the IGM generated by them would be smaller and less detectable, so that an alternative observing strategy than the one discussed here would be required.

We acknowledge helpful conversations on the subject of this paper with P. Palmer, L. Spitzer, M. Voit, W. Watson, and M. White. We thank the referee, G. Swarup, for valuable comments

which helped to clarify the detectability of a neutral IGM with the GMRT. P.M. and A.M. were supported in part by NSF grant PHY94-07194 to the ITP, University of California, Santa Barbara. A.M. is also grateful to the W. Gaertner fund for its support. M.R. thanks the Royal Society for support.

A. Ly α Line Flux in the Neutral Medium

We evaluate the Ly α flux produced in the neutral H I outside an ionization front from recombinations within the H II region itself. The volume emissivity is given by

$$\epsilon_\nu = x_e x_p n_{\text{H}}^2 \alpha_{2^2P}^{\text{eff}} h\nu \phi(\nu), \quad (\text{A1})$$

where $\alpha_{2^2P}^{\text{eff}}$ is the effective recombination coefficient to the 2^2P level (Pengelly 1964). Integrating from just outside the ionization front into the H II region, and averaging over 4π sr, we get

$$J_\alpha(r) = \frac{1}{4\pi} \int_{r_I}^r \epsilon_\nu dr \approx 1.2 \alpha_{2^2P}^{\text{eff}} \frac{n_{\text{H}}^2 hc}{8\pi H} \quad (\text{A2})$$

(with $x_e = 1.2$, $x_p = 1$, and including an additional factor of 1/2 since Ly α photons arise only from within the H II region). In evaluating the integral, we relate the distance $r - r_I$ from the ionization front to the frequency shift y from line center (in units of the thermal width) due to the Hubble expansion, by $r - r_I = by/H$, where b is the Doppler parameter. Because emission at line center is quickly redshifted out of the core, only the surface of the H II region contributes to the Ly α flux. The resulting intensity can be expressed as

$$J_{\alpha,-21}(z) \approx 0.03 h_{50}^{-1} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^2 \left(\frac{1+z}{7} \right)^{4.5}. \quad (\text{A3})$$

The recombination photons escaping from an expanding H II region produce then only a small flux of Ly α photons.

For completeness, we shall also compute the contribution from the local production of Ly α photons in the warm (mostly) neutral material surrounding the ionized gas. As Ly α photons are redshifted out of the Doppler core by the Hubble expansion, the local

background intensity becomes $J_\alpha = (n_{\text{H}}^2 hc / 4\pi H) [x_e x_p \alpha_{22P}^{\text{eff}} + x_e x_{\text{HI}} \gamma_{\text{eH}} + \mathcal{E} \dot{E}_X / (n_{\text{H}} h \nu_\alpha)]$, where $\gamma_{\text{eH}} \approx (2.2 \times 10^{-8} \text{ cm}^3 \text{ s}^{-1}) \exp(-11.84/T_4)$ (Osterbrock 1989) is the collisional excitation rate of Ly α photons by electron impacts, and the last term accounts for Ly α collisional excitations from the nonthermal distribution of electrons generated by the soft X-rays. The factor $\mathcal{E} \approx 0.4$ (Shull & van Steenberg 1985) represents the net fraction of photon energy converted to Ly α photons. For a temperature of $T_4 = 0.1$ and allowing (optimistically) for $x_e \sim x_p \sim 0.2$, radiative recombinations dominate, and

$$J_{\alpha,-21}(z) \approx 0.5 x_e x_p h_{50}^{-1} \left(\frac{\Omega_{\text{IGM}} h_{50}^2}{0.05} \right)^2 \left(\frac{1+z}{7} \right)^{4.5}. \quad (\text{A4})$$

Recombinations and collisional excitations in a warm, largely neutral medium are therefore a negligible source of Ly α photons.

B. Heating by Resonant Scattering of Ly α Photons

We derive the rate of transfer of energy, equation (30), from the radiation field to the atoms through recoil using the equation of radiative transfer. Neglecting expansion and source terms, the rate of change of the number density of photons in a given frequency band n_ν for an isotropic radiation field was shown by Basko (1981) to be given in the diffusion approximation by

$$\frac{1}{c} \frac{\partial n_\nu}{\partial t} = \frac{1}{2} n_{\text{HI}} \Delta \nu_D \frac{\partial}{\partial \nu} \left[\sigma_\nu \left(\Delta \nu_D \frac{\partial n_\nu}{\partial \nu} + 2 f n_\nu \right) \right], \quad (\text{B1})$$

where n_{HI} is the number density of hydrogen atoms in the ground state, $f = h \nu_\alpha / m_{\text{H}} c b$ expresses the effect of recoil, and only terms of order v/c have been retained. This is only correct as long as $kT_K \ll h \nu_\alpha$. Note that, in a steady state, $n_\nu \propto \exp[-h(\nu - \nu_\alpha)/kT_K]$ near the resonant frequency, as was first demonstrated by Field (1959b).

The rate of change of energy in the radiation field due to recoil is then

$$\dot{u}_\alpha = h \nu_\alpha \dot{n}_\alpha = n_{\text{HI}} c \Delta \nu_D h \int_0^\infty d\nu \nu \frac{\partial}{\partial \nu} (\sigma_\nu f n_\nu). \quad (\text{B2})$$

Performing the integration by parts, noting that σ_ν is sharply peaked at the resonant Ly α frequency, using $n_\nu = (4\pi/c) J_\nu / h \nu$, and setting $\dot{E}_\alpha = -\dot{u}_\alpha / n_{\text{HI}}$, we recover equation (30) of the text. This

result is valid in the presence of expansion and source terms as well, as may be shown from the generalization of equation (B1) by Rybicki & Dell’Antonio (1994). We note that, to order v/c , the flow of energy is one-directional, from the radiation field to the atoms. Treating the reverse flow requires the inclusion of terms of order $(v/c)^2$ (including relativistic effects like aberration). We have also neglected stimulated emission. These effects become particularly important when the matter approaches thermal equilibrium with the radiation field. In this limit, the radiation field relaxes to a blackbody spectrum, and the net transfer of energy ceases. In the situations relevant to this paper, these effects are negligible.

REFERENCES

- Adams, T. F. 1971, *ApJ*, 168, 575
- Allison, A. C., & Dalgarno, A. 1969, *ApJ*, 158, 423
- Allison, A. C., & Smith, F. J. 1971, *Atomic Data*, 3, 317
- Arons, J., & Wingert, D. W. 1972, *ApJ*, 177, 1
- Bahcall, J. N., & Ekers, R. D. 1969, *ApJ*, 157, 1055
- Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, *ApJ*, 304, 15
- Basko, M. M. 1981, *Astrophysics*, 17, 69
- Bebbington, D. H. O. 1986, *MNRAS*, 218, 577
- Bennett, C. L., et al. 1994, *ApJ*, 434, 587
- Bergeron, J., & Souffrin, S. 1971, *A&A*, 11, 40
- Bond, J. R. 1995, in *Cosmology and Large Scale Structure*, ed. R. Schaeffer (Netherlands: Elsevier Science Publishers), in press
- Braun, R. 1996, in *Cold Gas at High Redshift*, eds. M. Bremer et al. (Kluwer: Dordrecht), in press
- Couchman, H. M. P. 1985, *MNRAS*, 214, 137
- Couchman, H. M. P., & Rees, M. J. 1986, *MNRAS*, 221, 53
- Deguchi, S., & Watson, W. D. 1985, *ApJ*, 290, 578
- Fabbiano, G. 1989, *ARA&A*, 27, 87
- Field, G. B. 1958, *Proc. I.R.E.*, 46, 240
- . 1959a, *ApJ*, 129, 536
- . 1959b, *ApJ*, 129, 551
- Gallego, J., Zamorano, J., Aragón-Salamanca, A., & Rego, M. 1995, *ApJ*, 455, L1
- Górski, K. M., Ratra, B., Sugiyama, N., & Banday, A. J. 1995, *ApJ*, 444, L65

- Gould, R. J., & Thakur, R. K. 1970, *Ann. of Phys.*, 61, 351
- Gunn, J. E., & Peterson, B. A. 1965, *ApJ*, 142, 1633
- Hogan, C. J., & Rees, M. J. 1979, *MNRAS*, 188, 791
- Hu, W., & White, M. 1996, *A&A*, in press
- Longair, M. S. 1995, in *Extragalactic Background Radiation*, eds. D. Calzetti, M. Livio, & P. Madau (Cambridge: Cambridge University Press), p. 223
- Madau, P., & Meiksin, A. 1991, *ApJ*, 374, 6
- . 1994, *ApJ*, 433, L53
- Madau, P., & Shull, J. M. 1996, *ApJ*, 457, 551
- Mather, J. C., et al. 1994, *ApJ*, 420, 439
- Meiksin, A. 1994, *ApJ*, 431, 109
- Meiksin, A., & Madau, P. 1993, *ApJ*, 412, 34
- Miralda-Escudé, J., & Rees, M. J. 1994, *MNRAS*, 266, 343
- Moiseiwitsch, B. L. 1962, in *Atomic and Molecular Processes*, ed. D. R. Bates (New York: Academic Press), p. 280
- Osterbrock, D. E. 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei* (Mill Valley: University Science Books)
- Peebles, P. J. E. 1993, *Principles of Physical Cosmology*, (Princeton: Princeton University Press), Ch. 6
- Pengelly, R. M. 1964, *MNRAS*, 127, 145
- Phinney, E. S. 1985, *Astrophysics of Active Galaxies and QSOs*, ed. J. S. Miller (Oxford: Oxford University Press), p. 453
- Press, W. H., & Schechter, P. 1974, *ApJ*, 187, 425
- Purcell, E. M., & Field, G. B. 1956, *ApJ*, 124, 542

- Rees, M. J. 1985, MNRAS, 213, 75P
- Rybicki, G. B., & Dell’ Antonio, I. P. 1994, ApJ, 427, 603
- Schneider, D. P., Schmidt, M., & Gunn, J. E. 1991, AJ, 101, 2004
- Scott, D., & Rees, M. J. 1990, MNRAS, 247, 510
- Shapiro, P. R. 1986, PASP, 98, 1014
- Shull, J. M., & van Steenberg, M. E. 1985, ApJ, 298, 268
- Spitzer, L. 1962, Diffuse Matter in Space, (New York: Wiley-Interscience)
- Spitzer, L. 1978, Physical Processes in the Interstellar Medium, (New York: Wiley-Interscience)
- Spitzer, L., & Scott, E. H. 1969, ApJ, 158, 161
- Subramanian, K., & Padmanabhan, T. 1993, MNRAS, 265, 101
- Sugiyama, N., Silk, J., & Vittorio, N. 1993, ApJ, 419, L1
- Swarup, G. 1996, in Cold Gas at High Redshift, eds. M. Bremer et al. (Kluwer: Dordrecht), in press
- Swarup, G., Ananthakrishnan, S., Kapahi, V. K., Rao, A. P., Subrahmanya, C. R., & Kulkarni, V. K. 1991, Curr. Sci., 60, 95
- Swarup, G., & Malik, R. K. 1996, private communication
- Urbaniak, J. J., & Wolfe, A. M. 1981, ApJ, 244, 406
- Warren, S. J., Hewett, P. C., & Osmer, P. S. 1994, ApJ, 421, 412
- White, S. D. M., Efstathiou, G., & Frenk, C. S. 1993, MNRAS, 262, 1023
- Wouthuysen, S. A. 1952, AJ, 57, 31

Fig. 1.— Numerical calculation of the IGM surrounding a QSO (assumes $q_0 = 0.5$, $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\text{IGM}} = 0.05$). The QSO was turned on gradually at $z = 6$ with a specific luminosity of $L_\nu = 10^{31} \text{ ergs s}^{-1} H z^{-1} (\nu/\nu_L)^{-1.5}$. The effects of secondary electrons are significant for soft X-rays, and have been included, along with radiative cooling mechanisms and Compton cooling off the CBR. (a) Shows the temperature of the IGM due to soft X-ray heating by the QSO. At $z = 5.5$, the IGM is heated to a temperature above that of the cosmic background radiation out to a radius of 23 Mpc. The light front is at 25 Mpc. (b) Shows the H II (*solid*) and He II (*dashed*) fractions. The ionized region surrounding the QSO extends to 7.5 Mpc at this time.

Fig. 2.— Heating rate per hydrogen atom of the IGM due to soft X-rays from collapsed galaxy halos, in a flat $q_0 = 0.5$ and $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ CDM-dominated cosmology (*solid*). Also shown is the required heat input over a Hubble time to match the CBR temperature, $1.5T_{\text{CBR}}(z)/H(z)$ (*dotted*). For $z < 7$, the rate of X-ray heating from collapsed halos is sufficient to raise the temperature of the IGM above that of the CBR. At higher redshifts, the IGM is too dense and the hot halos that can produce the required penetrating soft X-rays are too few for efficient heating. At these epochs, only part of the IGM will be heated above the CBR temperature. The heating rate \dot{E}_{th} due to Ly α resonant scattering (*dashed*) is shown for a background radiation field sufficiently strong to couple the spin temperature to the kinetic temperature of the IGM (see text).

Fig. 3.— (a) The 21-cm efficiency is shown as a function of distance from a QSO with $L_{\alpha,47} = 1$ and $\alpha_S = 1.5$, turning on at $z = 8.7$. A Ly α radiation field of intensity $J_{\alpha,-21} = 0.8[(1+z)/9]$ was turned on at $z = 9$ and partially couples the spin temperature of the neutral hydrogen to the kinetic temperature. The effect of heating of the IGM by Ly α collisions is included. The efficiency is shown at time intervals of 2.5, 5, and 10×10^7 yr after QSO turn-on, corresponding to $z = 8.3$, 8.0, and 7.4, respectively. Because the IGM temperature is less than that of the cosmic background radiation at early times, a large efficiency is achieved in an X-ray warmed shell surrounding the quasar. (b) A map of the differential antenna temperature (K) as a function of angle (degrees) and velocity shift (10^3 km s^{-1}) (differenced from 21-cm redshifted to 153 MHz), resulting from two quasars turning on at $z = 8.7$, as in (a). One quasar is at a comoving position corresponding to

$z = 7.4$ and the other at $z = 8.0$. The quasars are separated by 2.5° on the sky. Such a map could be constructed by making a sequence of angular maps in a given frequency band, here taken to be 2000 km s^{-1} wide, and stepping in frequency until the full emitting region surrounding the quasars is recovered. In practice, a bandwidth of $5000 - 10000 \text{ km s}^{-1}$ will give a more feasible integration time for establishing a positive detection. The emission is measured relative to a reference beam well-separated from the quasars, either by frequency or by angle.







